Do property rights in personal data make sense after the big data turn

Purtova, Nadezhda

Published in:
Journal of Law and Economic Regulation

Publication date:
2017

Document Version
Peer reviewed version

경제규제와 법 (Journal of Law & Economic Regulation)
제10권 제2호 (통권 제20호). (Vol. 10. No. 2). 2017. 11. pp.0~0

Special Issue 영문★★

Do property rights in personal data make sense after the Big Data turn?†

Individual control and transparency

Nadezhda Purtova*

Contents

〈ABSTRACT〉

This paper offers an update – from the European perspective – to the debate on property rights in personal data. It argues that recent developments in the data processing technology and practices, specifically, the AI-driven Big Data Analytics, have rendered personal data a difficult object of enforceable individual property rights. There are two main reasons for this. First, data processing resulting from a decision of one person will inevitably have spill-over effects on others, e.g. as a result of profiling, or as a result of the same piece of data relating to a group of people, e.g. genetic data. This phenomenon is also called 'network effects'. Therefore, true individual control over personal data and also the effective enforcement of the individual property rights in personal data are difficult if not impossible to achieve. Second, creating and managing property rights that are transparent in terms of the object of property and the rights-holders is also challenging. This is due to the dynamic approach to the definition of personal data adopted in Europe: the same piece of data, depending on a particular context, can be personal and non-personal, more or less likely to relate to an identifiable natural person, and with a stronger or weaker link to that person. While this does not necessarily constitute a problem for the purposes of the data protection law, and the broadest definition of personal data can achieve the goals of complete and effective protection, enforcing property rights in personal data is difficult. The difficulty lies, first, in determining at which point the level of relation to an individual is sufficient to establish property rights, and second, in tracing the presence of such a relation.

† 투고일자 2017. 00. 00, 심사일자 2017. 00. 00, 게재확정일자 2017. 00. 00.

* Dr Nadezhda Purtova is Associate Professor at Tilburg Institute for Law, Technology, and Society, the Netherlands n.n.purtova@uvt.nl

Keywords: Property in personal data, data ownership, transparency of property rights, collective rights in personal data, data commons, data as a common good, data portability, data as counter-performance, data as an asset, access to data

Ⅰ. Introduction

In 2016 Klaus Schwab, the founder of the World Economic Forum, coined the term 'Fourth Industrial Revolution' to refer to the ongoing technology-led transformation of how we live and work that has been unfolding across nations and industries. According to Schwab, the Fourth Industrial Revolution builds on the advances in information technology and the resulting digitization, or 'the Third Industrial Revolution'. Yet, it fundamentally differs in 'velocity, scope and systems impact', leading to the 'fusion of technologies that is blurring the lines between the physical, digital, and biological spheres'.[1] Unprecedented connectivity of people and devices embodied in the Internet of Things ('IoT') and smart environments, advances in Artificial Intelligence ('AI') merging with robotics are just some of the forerunners of the Fourth Industrial Revolution.

Data is the lifeblood of this transformation. To name just a few examples, data is generated by and transferred through the connectivity infrastructures; the AI algorithms 'feed' on it and learn to better recognize relevant patterns. Data has become an essential part of (knowledge) production, management and governance; think of the data-driven production of digital content and services, data-driven (precision) medicine, data-driven process control, e.g. in agriculture, prediction and decision-making in private and public sectors. Underlying it all, data has been acknowledged as an essential resource for economic growth. It is estimated that in Europe alone by 2020 the size of the data economy may increase to €739 billion, or 4% of the overall EU GDP.[2] Therefore, the question of data ownership and the related questions of data access are key to maintaining a degree of control over this technological, economic and societal transformation that the Fourth Industrial Revolution is said to be.

As the Call for Contributions invited to examine possible regulatory paradigms that can 'lead the ongoing Fourth Industrial Revolution', this paper will focus on the property rights in data as one such regulatory paradigm. More specifically, this paper will address property rights in personal data, since a significant and still growing share of data relates to identified or identifiable natural persons, and hence constitutes 'personal data' in the sense of the EU data protection law.

The question this paper will answer at least in part is if individual property rights in personal data still make sense in Europe. The analysis will rely on the European policy and legislative framework, inter alia, regarding the meaning of property rights and the goals and content of the data protection law. Yet, the arguments in this paper could also be used outside Europe in the same contexts where the European propertisation debate takes place[3] and to the extent that the general-

1) K Schwab 'The Fourth Industrial Revolution: what it means, how to respond' (2016) www.weforum.org, published 16 January 2016.

2) European Commission 'Building a European Data Economy', published 10 January 2017, available online at <https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy>, last accessed 27 September 2017.

3) See 2 for description of the contexts.

izations about the meaning of property are correct.

While the matter of property in personal data has received considerable attention in the literature since 1970s, the discussion is in need of a serious update. Indeed, while little new analyses were added to the property in personal data literature in the past years,[4] data processing technologies and practices as well as the European policy and legislative landscape have made giant leaps that need to be taken into account.

Specifically, the adoption in April 2016 of the General Data Protection Regulation[5] ('GDPR') marked the end of the European data protection reform and is a major development on a legislative level. According to the European Commission, one of the general objectives of the reform was "[t]o increase the effectiveness of the fundamental right to data protection", which implied, among others, "that individuals are in control of their personal data and trust the digital environment."[6] As a result, the GDPR contains new rights considered by some to be property-like, e.g. the rights to data portability and to erasure ('the right to be forgotten').[7]

The continued datafication and advances in the AI-enabled Big Data Analytics are another major change that needs to be considered in the property debate. Namely, datafication stands for the modern data mining and advanced data analytics techniques that are able to turn *literally everything* into data, including what we never thought of as containing or being data at all.[8] In addition, the AI-enabled Big Data Analytics is able to lay connections between people and data quicker and differently from how it was done before. These phenomena together result in the general increase in the quantity of data available and more data being 'personal'. This fundamentally alters the scale of the debate on property rights in personal data, the change being both quantitative and qualitative.

Before the analysis can begin, I will briefly deal with the matter of nomenclature. In the context of this analysis a property right in law is understood as any legally protected interest in an object, tangible or intangible, that is directed against the entire world and hence has a so-called *erga omnes* effect.[9] This meaning of property rights is not attached to any one jurisdiction, but derives from studies in comparative European property law. Property rights understood as the rights to exclude, alienability, or the ability to sell, is ~~therefore~~ not a necessary defining characteristic of property.[10] To be recognised as inter-

4) Most analyses that *are* published in the part 5 years are relying on the propertisation arguments developed before 2012 and hence pre-Big Data (e.g. S Spiekermann, A Acquisti, R Böhme and KL Hui 'The challenges of personal data markets and privacy' (2015) 25(2) *Electronic Markets*, 161-167).

5) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), published in *Official Journal of the European Union*, Vol 59, 4 May 2016.

6) Table 1 in European Commission, 'Impact Assessment Accompanying the document Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) and Directive of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data by competent authorities for the purposes of prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and the free movement of such data' Commission Staff Working Paper SEC (2012) 72 final, Brussels, 25.1.2012, 43.

7) This point is further developed in 2.1.

8) V Mayer-Schönberger and K Cukier, *Big data: A revolution that will transform how we live, work, and think* (Houghton Mifflin Harcourt 2013) 29.

9) N Purtova, *Property Rights in Personal Data: a European Perspective* (Kluwer Law International 2011) 57 et seq.

10) ibid. 86-88. But see E Ostrom and C Hess, 'Private and Common Property Rights' (2010) in B Bouckaert (ed.)

ests of a proprietary nature, and to consequently enjoy proprietary status, rights must comply with two 'leading principles of property law' or 'filters', as Van Erp describes them:[11] the *numerus clausus* principle, i.e. the rights have to be on the list of property rights recognised as such by law, and their content cannot be modified at all, or can only be modified a little; and the principle of transparency, i.e. the rights have to be made public, either by registration (in the case of immovable property) or possession (in the case of movable objects).[12]

The argument will progress in the following steps. Part 2 will provide some background of the propertization debate by introducing three contexts in which the property in personal data talk is relevant and arguably makes sense: (1) privacy-protective potential of property rights, (2) economic interest and investment in data, and (3) data access. Importantly, this paper will not advocate in favor of introducing property rights in personal data for any of the three reasons, but rather list and elaborate on these reasons as plausible rationales of propertization that deserve consideration. Part 3 is where the core of the argument will lie. It will explore the feasibility of introducing and managing property rights in personal data from the perspective of individual control in personal data and transparency of property rights if introduced. Part 4 will conclude the paper with a discussion.

Property Law and Economics (Edward Elgar Publishing 2010) 338: 'Property-rights systems that do not contain the right of alienation are considered to be ill-defined.'.

11) S Van Erp, 'From "Classical" To Modern European Property Law?' *Essays in honour of Konstantinos D. Kerameus* (Bruylant 2009) Available at SSRN: https://ssrn.com/abstract=1372166, 10.

12) *Ibid.*

Ⅱ. Context: Three Reasons to Talk about Property in Personal Data

1. Privacy-protective potential of property rights. Individual control over ~~their~~ data

Proposals to introduce property rights in personal data have emerged in the United States as early as the 1970s,[13] and have been subject of academic discussion ever since. Property ~~right~~ understood as the right to exclude, a significant part of the debate considered propertization as a means of giving back to the individual control over data pertaining to him or her.[14] Some arguments have been made against propertization, a predominant anti-propertization argument being that informational privacy is a public good and propertization facilitating market exchange would not be able to secure it.[15] In response, other scholars have offered property models consistent with and ~~arguably~~ enhancing informational privacy.[16]

In Europe the idea to introduce property rights in personal data to achieve data protection goals gained traction since property was considered by some particularly suitable to secure the core of the European data protection, i.e. individual control over

13) A. Westin, *Privacy and Freedom* (Atheneum 1967).

14) e.g. E J Janger, 'Muddy Property: Generating and Protecting Information Privacy Norms in Bankruptcy' (2003) 44 WM. & MARY L. REV. 1801; J Rule, *Privacy in Peril: How We Are Sacrificing a Fundamental Right in Exchange for Security and Convenience* (Oxford: Oxford University Press 2007).

15) e.g. P Regan, 'Privacy as a common good in the digital world' (2002) 5(3) *Information, Communication & Society,* 382.

16) P Schwartz, 'Privacy, property and personal data' (2004) 117 *Harvard Law Review* 2056; J Rule (n 14); L Lessig, *Code 2.0* (New York 2006); E J Janger (n 14), etc. See Chapter 6 in N Purtova (n 6) for a more detailed account of the various arguments against and in favor of property rights in personal data.

one's data and informational self-determination. Yet, at the time of writing, under EU data protection law, property rights in personal data are not acknowledged explicitly and remain ill-defined.[17]

As I argued elsewhere,[18] what makes property rights a suitable legal instrument to achieve individual control over one's personal data is its *erga omnes* effect, enabling data subjects to enforce their control over personal data against the entire world. This is especially useful in the complex conditions of the modern data flows where location of data and the chain of control over it are often hard to trace to known contract parties. A system of personal data licenses based on the default control rights of the data subjects – akin to what Paul Schwartz, James Rule and Edward Janger propose in the context of the United States[19] - would create a coherent and more articulate framework for personal data management that allows data use but also is respectful of the principle of information self-determination,[20] where information self-determination is understood as 'the capacity of the individual to determine in principle the disclosure and use of his/her personal data'.[21]

To achieve greater insight into how the property regime could grasp the complexity of the modern relationships vis-à-vis personal data, and form a regulatory framework for the data flow that is respectful of the information self-determination, it is helpful to look at the system of English land law. Briefly, English land law governs what a continental lawyer would call 'property rights in immovables'. Like personal data, land is a valuable resource that is transferred to multiple actors, who put it to many uses.[22] To accommodate these uses, and grant protection to respective interests in land, modern land law developed into a pyramid-like system of rights and interests, with the right with the most broad scope - fee simple - at the bottom, and leases - property rights of a narrower scope - at the top. The content of these rights has been tailored to account for the most popular uses of land, and, according to the principle of *numerus clausus*, no other rights in land, save for those on the list of recognized property entitlements, receive *erga omnes* protection. The transfer of leases - the 'lesser' rights in a piece of land - does not undermine, although it does limit, the 'greater' right of fee simple. However, at all times, until the fee simple is transferred in full, its holder retains some control over his property, e.g. the right of access in order to maintain an object of property rights in a proper state, etc.

In a search for this quality, namely the capacity to exercise control over a transfer and retain some control after the transfer takes place, a similar system of property rights could be built around personal data. An individual - the data subject - could be said to have the broadest property right possible (although it would not be unlimited), including a right to transfer his or her personal data for remuneration. The most important limitation

17) As argued in N Purtova 'Default entitlements in personal data in the Proposed Regulation: Informational Self-Determination Off the Table … and Back On Again?' (2014) 30(1) *Computer Law and Security Review,* 6.

18) See Purtova (n 16) for a full argument on how property rights in personal data can achieve data protection purposes.

19) Purtova (n 17).

20) Purtova (n 16).

21) P De Hert and S Gutwirth, 'Data protection in the case law of the Strasbourg and Luxemburg: Constitutionalisation in action' in S Gutwirth et al (eds) *Reinventing data protection* (Springer 2009) 14.

22) Purtova (n 16) in Chapter 4 demonstrates that in modern property law whether an object is physical, like land, or intangible, like personal data, does not matter.

on the possible scope of this right would be a prohibition on the waiver of the data protection guarantees.[23] In Europe where data protection has been recognized as a human and fundamental right,[24] the core limitation on data transfers would be that a data subject may not waive the human rights guarantees, and so cannot relinquish the control over his or her personal data entirely.[25] On this basis, it would not be possible for the control rights in personal data to be completely alienated for remuneration or counter performance or waived.

The smaller rights in personal data that would be transferred from a data subject are comparable to leases in land law; the alienable 'leases' in personal data could be tailored to reflect most common uses of personal data, and could also vary in type, depending, e.g. on the duration and purpose limitations, e.g. excluding the use of the data for profiling. Moreover, pursuant to the principle of *numerus clausus*, recognising only a closed list of 'lesser' property rights in personal data would be one step further along the road to ensuring that individuals are not forced into relinquishing total control over their personal information.

It would be a matter of policy as to whether actors other than a data subject should be permitted to enjoy property rights over personal data, or whether a situation in which the individual is the only holder of property rights over his personal information should be maintained. ~~In the latter case scenario the exercise of transfers from one actor to another will be on the basis of a contract.~~

Such a system of property rights could, in theory, be implemented through the use of so-called 'sticky technologies' which enable the rights relating to a piece of data to 'travel' with it and help verify compliance of a particular data processing operation with the conditions imposed on it.[26] In sum, the protective potential of the individual property rights in personal data is in providing an alternative and arguably more effective legal tool facilitating individual control over collection and use of the data pertaining to the individual.

2. Acknowledging economic interest and investment in data

As a part of its Digital Single Market Strategy, Europe has declared its commitment to develop as a data economy.[27] Data has been acknowledged as an essential resource for economic growth, and it is estimated that by 2020 the size of the EU data economy may increase to €739 billion, or 4% of the overall EU GDP.[28] Against this back-

23) N Purtova 'Private law solutions in European data protection' (2010) 28(2) *Netherlands Quarterly of Human Rights* 179.

24) EU Charter of Fundamental Rights in Art. 8, and most recently, in *Satamedia* case where the ECtHR seems to have expanded the scope of Art 8 ECHR to cover data protection regardless of the kind of data processed (~~paae~~ 137 of the judgment reads as follows: "Article 8 of the Convention thus provides for the right to a form of informational self-determination, allowing individuals to rely on their right to privacy as regards data which, albeit neutral, are collected, processed and disseminated collectively and in such a form or manner that their Article 8 rights may be engaged.").

25) Purtova (n 16), Chapter 9.

26) E.g., see J Cohen, 'Examined Lives: Informational Privacy and the Subject as Object' (2000) 52 *Stanford Law Review* 1391.

27) European Commission, "Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions. A Digital Single Market Strategy for Europe" COM (2015) 192 final, 6 may 2015, Brussels, 14.

28) European Commission 'Building a European Data Economy', published 10 January 2017, available online at <https://ec.europa.eu/digital-single-market/en/policies/building-europ

ground, legal rules on allocation and extent of control over data, i.e. exclusive rights and the rights of access, become increasingly important. Therefore it comes as no surprise that the EU policymakers intend to address the issue of data ownership.[29] Recent EU law addresses both individuals' and companies' economic stakes in personal data through property-like rights.

On the individual level, GDPR introduced the right to data portability in relation to personal data. A number of scholars suggest that data portability is conducive of or closely akin to the property-rights approach to data protection, or data ownership.[30] Albeit, these authors seem to focus on what Rubinstein calls "property-related actions like trading, exchanging, or selling data", rather than the defining element of property rights, i.e. the right to exclude. Under Article 20 GDPR, an individual to whom the data relates ('data subject') has a right to receive a copy of personal data pertaining to him or her in a structured, commonly used and machine-readable format and to transmit those data to - in the data protection parlance - another 'controller', i.e. any person or legal entity who determines the purposes and means of data processing. According to Article 29 Working Party, the EU advisory authority on data protection, '[t]he primary aim of data portability is enhancing individual's control … and making sure they *play an active role in the data ecosystem*',[31] among others, by preventing service lock-ins. Tene and Polonetsky consider data portability in a reusable format a way for the information industry to share the wealth created from personal data with those individuals, and predict that it will create a market for the user-centric personal-data applications providing the data analytics and management services for the benefit of the data subjects alone.[32]

Although not relating to the right to exclude, the proposed Digital Content Directive[33] acknowledges the economic role of personal data in digital economy as asset or even currency, among others, as counter-performance by the consumer in exchange for digital content:

> *In the digital economy, digital content is often supplied without the payment of a price and suppliers use the consumer's personal data they have access to in the context of the supply of the digital content or digital service.*[34]

The Directive will apply to such contracts (Article 3). Among others, in the event of termination of a contract, the supplier will bear an obligation to reimburse fully or partially payment under the contract (Article 13a(1) and (2)) and comply with the Article 20 GDPR data portability obligations when personal data is concerned (Article 13a(3)).[35]

ean-data-economy>, last accessed 27 September 2017.

29) European Commission (n 27).

30) e.g. IS Rubinstein, 'Big Data: e End of Privacy or a New Beginning?' (2013) 3(2) *International Data Privacy Law*, 74; P Swire & Y Lagos, 'Why the right to data portability likely reduces consumer welfare: antitrust and privacy critique' (2013) 72(2) *Maryland Law Rev.* 335, 373.

31) Article 29 Working Party, 'Guidelines on the right to data portability', 5 April 2017, 16/EN WP 242 rev.01., 4, fn 1.

32) O Tene and J Polonetsky, 'Big data for All: Privacy and User Control in the Age of Analytics' (2013) 11(5) *Northwestern Journal of Technology and Intellectual Property*, 264.

33) EU Presidency, *General approach on Proposal for a Directive of the European Parliament and of the Council on certain aspects concerning contracts for the supply of digital content first reading)*, published in Brussels on 1 June 2017.

34) Ibid., 8, fn 15.

35) For a more detailed analysis of the right to data portability under GDPR and outside of the data protection, including the proposed Directive on digital content, see I De Graef, M

At the same time, the EU Intellectual Property law acknowledges quasi-property claims of the information industry in personal data. In particular, personal data aggregated by the information industry may be subject to *sui generis* database right or be protected as a trade secret.[36] *Sui generis* database rights protect exclusive rights in databases resulting from 'substantial investment' in the collection, verification or presentation of data;[37] the 'substantial investment' expressed, e.g. in efforts of collection, classification or cleaning. Protection is granted against extraction or reuse of substantial part of the database, or of unsubstantial part in case of extraction and reuse which are systematic (Article 7 Database directive). Trade secrets protect commercial information where its economic value to a firm hinges on it remaining secret,[38] for instance, customer lists and profiles. The protection is granted against unlawful acquisition of secrets (Article 4(2) of the Trade secrets directive). It is predicted that the IP claims in personal data and the Article 20 GDPR data portability right will clash once the GDPR enters into effect.[39]

3. Sharing in data-generated wealth and data access: data 'haves' and 'have-nots'

Given the high economic value of personal data it is no surprise that the generated wealth is not distributed equally, when the large digital platforms capitalize on personal data of their users, among others, through rapid expansion of intellectual property rights,[40] and the users see none of it. This is akin to what Boyle describes as the "Second Enclosure Movement" in the context of knowledge and ideas, or as he puts it, the 'intangible commons of the mind'.[41]

To illustrate, Evgeny Morozov in his *Financial Times* piece[42] has argued that digital giants like Google and Facebook harvest, hoard, hold monopoly over and exclusively profit from the pools of data collected through their various services, whereas these pools are not available to anyone else. Alternatively, the large private platform providers are in full control of access to data which increasingly becomes a critical resource. Everyone else: people, researchers, smaller data-centred businesses and others are precluded from profiting from these valuable data assets, unless granted access by the platform providers. The latter thus become the gatekeepers of data and access to the data-generated forms of knowledge.

I have argued elsewhere[43] that the lack of clear allocation of ~~de jure~~ property rights to the

Husovec and N Purtova 'Data Portability and Data Control: Lessons for an Emerging Concept in EU Law' (forthcoming).

36) The following overview of the IP claims in data substantially relies in the overview in I De Graef, M Husovec and N Purtova, ibid.

37) Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases ('Database Directive') [1996] OJ L 77/20 and Case C-203/02, *The British Horseracing Board*, EU:C:2004:695.

38) Directive 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure ('Trade Secret Directive') [2016] OJ L 157/1.

39) I De Graef, M Husovec and N Purtova, n 35.

40) See 2.2.

41) J Boyle, 'The Second Enclosure Movement and the Construction of the Public Domain' (2003) 66(1-2) *Law and Contemporary Problems* 33-74.

42) E Morozov 'Europe is wrong to take a sledgehammer to Big Google' *The Financial Times,* 12 January 2015 available online at www.ft.com.

43) N Purtova, 'The illusion of personal data as no one's property' (2015) 7(1) *Law, Innovation and Technology,* 83-111.

individuals by default is conducive of this data enclosure, where a few actors within the personal data-intensive Information Industry with the strongest market power are able to make and effectuate the strongest de facto exclusive claim on personal data, resulting in the division between the data 'haves versus the have-nots, the elite versus the masses.'[44] The divide may lie between the large private digital platforms and everyone else, the rich and the poor, as well as along the boundaries of the 'first', 'second' and 'third' worlds, race and social class.[45] Allocating property rights in personal data to individuals and effectuating them through, e.g. the data portability instruments,[46] will arguably help avoid this enclosure at least in part within the European borders.

Ⅲ. Feasibility of Property Rights In Personal Data Revisited

1. Individual control and network effects of personal data choices

As ~~Section 2~~ shows, from the data privacy perspective, the main benefit behind of introducing property rights in personal data is creation of an effective legal tool to enforce individual control over his or her personal data. Property rights language is arguably just another way for the data protection rights to be phrased. Yet, the idea that an individual can in fact exercise the control rights effectively has been criticised as naïve for a while now. The primary example of such criticism is the discussion around consent as a ground of legitimate data processing in Europe.[47] This paper will not engage with this discussion further, but only highlight its core. In short, leaving aside the debate on (the absence of real) consent in 'take it or leave it' situations and the accessibility of the privacy policies to a lay person, which could be helped with regulation and enforcement, there is a number of substantive points where the concept of consent seems to fail without major reshaping.[48] In the age of constant data collection and hundreds of data processing operations pertaining to one individual each day, it is believed that it is too much to ask of an individual to make truly informed decisions about each data processing operation, whether or not he or she wishes for ~~his/her~~ data to be processed. It is also a common argument that given the length and complexity of the privacy policies, sometimes reaching hundreds of pages and written in legalese, it is unreasonable to expect that an individual can read and comprehend them and give a truly informed consent. Due to the phenomenon called bounded rationality an individual is arguably unable to anticipate on the full range of consequences of ~~his/her~~ consent and the resulting data processing.[49] Finally, the field of behavioural economics of privacy has demon-

44) C Hess and E Ostrom 'Introduction: An overview of the knowledge commons' (2007) in C Hess and E Ostrom (eds.) *Understanding Knowledge as Commons* (MIT Press 2007) 13.

45) L Taylor, 'What Is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally', June 26, 2017, http://dx.doi.org/10.2139/ssrn.2918779.

46) See 2.2.

47) e.g. E Kosta *Consent in European data protection law* (Martinus Nijhof 2013).

48) The proposals for adjustment of consent include collective and assisted consent in L Bygrave and D Schartum, 'Consent, proportionality, and collective power' in *Reinventing data protection* (Springer 2009) and R. Brownsword 'Consent in data protection law: Privacy, fair processing and confidentiality' in S Gutwirth et al (eds) *Reinventing data protection* (Springer 2009).

49) E.g. RAND corporation, 'Review of the European Data Protection Directive: Technical Report Prepared for the Information Commissioner's Office', 29-30.

strated that individual control and consent are prone to manipulation, depending on how the conditions of consent are formulated.[50] Bygrave and Schartum and Brownsword suggest that with some adjustments in interpretation and application, consent and hence individual control should continue to be one of the key data protection rules, e.g. in the form of collective or assisted consent.

However, the ability of an individual alone to exercise control over his or her data is not in the centre of this analysis. The further discussion therefore will focus on whether or not such control should be *individual,* since the effects of data processing following the individual consent are often not limited to that individual alone. Most recent academic thinking on the ripple effects of privacy choices has been focusing on the potential of personal data originating from one person to impact others. Genetic data is ~~an~~ often used as an example. While this may be a decision of one to share ~~his/her~~ genetic data, the information revealed will pertain not only to the person sharing, but also to an entire group of ~~his/her~~ blood relatives over several past and future generations.[51] ~~Remarkably~~, in the age of proliferation of data collection any personal data, in a way, is like genetic data: one may decide to live a device-free life not to be subjected to decisions made on the basis of data processing. Yet, there will always be a significant group of people like ~~him / her~~ - e.g. neighbours sharing a postal area code, gender, age, or other characteristics - who are wiling to disclose their data, and modern analytics will be able to process more seemingly irrelevant data to produce information of significance that would contribute to a profile that would eventually be applied to the person who meant to opt out. For example, health-, socio-economic and other data extracted from a 'smart community' in Africa used essentially as a 'data farm' to develop a disease profile could also be applied to shape life of communities thousands of kilometres away, on other continents. For the reasons of these network effects of individual privacy choices and because in the context of the modern information practices no personal data remains strictly personal, there is a growing understanding of insufficiency of the individual data protection rights and a growing number of arguments in favour of group privacy[52] and some form of collective rights in data.[53]

2. Transparency of property rights and blurry boundaries of the concept 'personal data'

As explained in the Introduction, transparency is one of the two leading principles of property law that separates property- from non-property rights.[54] The principle of transparency dictates that in order to be recognized as property rights, the rights have to be made public, e.g. by registration (in the case of immovable property) or

50) L Brandimarte, A Acquisti and G Loewenstein, 'Misplaced Confidences: Privacy and the Control Paradox' (2012) 4(3) *Social Psychological and Personality Science*, 340–347.

51) e.g. BR Goldman, 'Pharmacogenomics: Privacy in the Era of Personalized Medicine' (2005) 4(1) *Northwestern Journal of Technology and Intellectual Property*, 84; J Gniady, 'Regulating Direct to Consumer Genetic Testing', 76(5) *Fordham Law Review*, 2429; Article 29 Working Party, 'Working Document on Genetic Data' Adopted on 17 March 2004, 12178/03/EN (WP 91), 4.

52) e.g. L Taylor, L Floridi and B van der Sloot (eds.) *Group privacy* (Springer 2017).

53) e.g. A Montelero 'Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection' (2016) 32(2) *CLSR,* 238–255.

54) Van Erp (n 11) p. 10.

possession (in the case of movable objects).[55] This is because the effect of property rights against everyone (as opposed to against contractual parties for contract-based rights) is the strongest and imposes the heaviest burden, and so, to comply, everyone needs to know when and in relation to which object property rights exist. This section will show that personal data is a difficult object of property rights when it comes to the principle of transparency, which has to do with the increasingly blurry boundaries of the concept 'personal data'.

The 1995 Data Protection Directive defines personal data as 'any information relating to an identified or identifiable natural person ('data subject')' (Article 2(a)). According to Recital 26, 'to determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person'. GDPR adopts the same definition (Article 4(1)) and Recital 26 of GDPR similarly requires that 'account should be taken of all the means reasonably likely to be used either by the controller or by any other person to identify the individual directly or indirectly.'

Article 29 Working Party ('WP29') adopted a non-binding opinion aimed to streamline the national implementation of the definition of personal data.[56] WP29 explains the definition of personal data as comprised of three essential elements: personal data is (1) any information that (2) relates to ~~people~~ who ~~are~~ (3) personally identified or identifiable, directly or indirectly.[57]

55) *Ibid.*
56) Article 29 Working Party (2007) Opinion 4/2007 on the concept of personal data, published 20 June 2007, (WP 136).
57) *Ibid.*

Let us begin with the identifiability. It is widely acknowledged that - given the condition of the modern data processing technologies, namely, growing technical ability to identify previously anonymous data sets and the resulting gradual failure of anonymisation, the same piece of data may be more or less easily identifiable to an individual. Famously, film rating records of 500,000 Netflix subscribers were re-identified in 2008 using the openly accessible Internet Movie Database.[58] In 2013 travel routes of celebrities such as Bradley Cooper and Olivia Munn, including street addresses, and whether or not they left a tip, where deduced from the "anonymised" public database of the New York taxi rides which contained no passenger information, and paparazzi pictures.[59] In 2014 knowing location of credit card holders on 4 occasions allowed to re-identify 90% of 3 months of credit card transactions, chronicling the spending of 1.1 million people in 10,000 shops, having access only to amounts spent, shop type and a code representing each person. Knowing the amounts spent on these 4 occasions lead to re-identification of nearly all card-holders.[60]

As the EU Court of Justice ruled recently in the *Breyer* case[61] data may be treated as data relating to an 'identifiable natural person' where the additional data necessary in order to identify are

58) A Narayanan and V Shmatikov, 'Robust de-anonymisation of large datasets. (How to break anonymity of Netflix Prize dataset)' (2008) *Proceedings - IEEE Symposium on Security and Privacy,* 111.
59) JK Trotter 'Public NYC Taxicab Database Lets You See How Celebrities Tip', *Gawker* 23 October 2014, available http://gawker.com/the-public-nyc-taxicab-database-that-accidentally-track-1646724546, accessed 25 August 2017.
60) J Bohannon, 'Credit card study blows holes in anonymity' (2015) 347 (6221) *Science*, 468.
61) Court of Justice of the European Union, Case C-582/14, *Patrick Breyer v. Bundesrepublik Deutschland*, judgment of 19 October 2016, ECLI:EU:C:2016:779.

held by a third party (among others, paras. 41 and 43). What matters is if combining the necessary data 'constitutes a means likely reasonably to be used to identify the data subject' (para. 45). It would not be the case 'if the identification of the data subject was prohibited by law or practically impossible on account of the fact that it requires a disproportionate effort in terms of time, cost and man-power, so that the risk of identification appears in reality to be insignificant' (para. 46). WP29 explains that, in addition, a number of other factors should be accounted for in determining the likelihood of identification: '[t]he intended purpose, the way the processing is structured, the advantage expected by the controller, the interests at stake for the individuals, as well as the risk of organisational dysfunctions (e.g. breaches of confidentiality duties) and technical failures.'[62] Crucially, Art 29 WP explains that 'the test is a dynamic one and should consider the state of the art in technology at the time of the processing and the possibilities for development during the period for which the data will be processed.'[63] While identification may not be possible at the time of processing, the state of art of technology, degree of data aggregation and hence the likelihood of identification may change, rendering the same data set identifiable.

This is exactly because of this fluid boundary between identifiable and non-identifiable data that establishing, exercising, and managing transparent property rights in personal data might prove problematic. Since the state of identifiability may change from low likelihood to 'reasonably likely', and the change may well take place unnoticed both by the data holder and data subjects, determining when (and whose) property rights emerge, what rights must be respected and can be enforced in relation to which data is one of the key challenges for propertization.

Let us now move to the first two elements of the definition of personal data which also present problems when in comes to the property rights transparency and clarity of the conceptual boundaries of personal data. In order to be considered personal data, in addition to being identifiable to an individual, data needs to relate to ~~or mean something about~~ that individual. This follows from the part of the definition where personal data is any information that *relates to* ~~people~~. According to WP29, for the purposes of the definition, information should be considered relating to ~~people~~ in a relevant way when it relates in content, purpose or result, i.e. respectively, when it is *about* that person (e.g. home address, list of university grades, or results of a medical test); or it is used or is likely to be used with the *purpose* to evaluate, treat in a certain way or influence status or behaviour of a person; or it has or is likely to have *impact* on a person's rights and interests, even in minor ways.[64] Some information is perceived as relevant more easily, for instance, information 'generated' *by* (observing) people (e.g. administrative records of people's off-line lives, and digital records of online behaviour like websites visited, texts and images uploaded; information generated through use of 'smart' objects and devices like phones or fitness bracelets), or objects people interact with (their cars, homes, computers). At the same time, some information is hard to intuitively place in any connection of relevance for anyone: e.g. the amount of weight a

62) Opinion 4/2007 (WP 136) of 20 June 2007, p. 15.
63) *Ibid.*
64) *Ibid.*, 9 et seq.

block of concrete can withstand, or the amount of sand crystals in a cubic meter of sand in the Sahara desert. In order to establish transparent property rights in personal data it is therefore important whether or not there is a link between an individual and a piece of data 'that matters.'[65] In the opinion on the *Nowak* case pending before the EU Court of Justice regarding status of an examination script as personal data, AG Kokkot suggests that the criterion of relation to an individual is also dynamic.[66] The same bit of data can be of no relation to an individual under some circumstances, and relate to that individual under another set of circumstances, likely depending on the similar factors as identifiabilty.

However, how meaning is 'attached' to data by modern machines is beyond the grasp of human mind.[67] The game-changer is a new generation of data-processing algorithms based on machine learning. Machine learning is the ability of computer algorithms to learn from data and make predictions for new situations,[68] and improve automatically through experience.[69] The new algorithms are autonomous, i.e. self-learning, self-repairing, and self-managing and form the core of the modern approach to Artificial Intelligence (AI), a strand of computer science aimed to build computers as intelligent agents. The way advanced AI self-learning algorithms make sense of data is not transparent even for their designers. Hence, the new AI algorithms work as a black box that is truly beyond human cognition.[70] These AI self-learning autonomous machines together with unprecedented amount of data already stored in databases or live-streamed, form the essence of Big Data[71] and have the ability to harness information in fundamentally novel ways.[72] In effect, we can no longer say that some data has no meaning. In fact, it is safer to assume that all data potentially has meaning, even if not for humans. This proposition may have a slight science fiction ring to it; yet, it will become more real as the AI-operated 'smart' communities and infrastructures become more and more embedded in our daily lives, where increasing number of parameters is quantified and processed by the AI algorithms in order to subject the users of intelligent infrastructures to the data-driven decisions. This blurring line between data that means something as opposed to data that means nothing with regard to an individual adds an additional level of complexity, first, in determining at

65) According to Art 29 WP, "this building block of the definition [relate to] is crucial as it is very important to precisely find out which are the relations/links that matter and how to distinguish them" (ibid.).

66) "It is true that the extent of the link between an examination candidate and his performance in an examination increases according to the extent to which he has to formulate the answers himself." (Case C-434/16 *Peter Nowak v Data Protection Commissioner*, ECLI:EU:C:2017:582, Opinion of Advocate General Kokott, delivered on 20 July 2017, para. 23).

67) M Hildebrandt, 'The Dawn of a Critical Transparency Right for the Profiling Era' in J. et al Bus (ed), *Digital Enlightenment Yearbook* (IOS Press 2012) 53; JP van Bendegem, 'Neat Algorithms in Messy Environments', in M. Hildebrandt and S. Gutwirth (eds.), *Profiling the European Citizen. Cross-Disciplinary Perspectives* (Springer 2008), 80-83.

68) J Stajic, R Stone, G Chin, and B Wible, 'Rise of the Machines' (2015) 349 (6245) *Science* 248-249 (published 17 July 2015).

69) MI Jordan and TM Mitchell, 'Machine learning: Trends, perspectives, and prospects' (2015) 349(6245) *Science*, published 17 July 2015, 255.

70) M Hildebrandt n 67, 53; JP Van Bendegem, 'Neat Algorithms in Messy Environments', in M. Hildebrandt and S. Gutwirth (eds.), *Profiling the European Citizen. Cross-Disciplinary Perspectives* (Springer 2008), 80-83.

71) M Hildebrandt, 'Slaves to Big Data. Or Are We?' (2013) 16 *IDP Revista De Internet, Derecho Y Política*, published December 2013, available online at http://works.bepress.com/mireille_hildebrandt/52/.

72) V Mayer-Schönberger and K Cukier, *Big data: A revolution that will transform how we live, work, and think* (Houghton Mifflin Harcourt 2013), 29.

which point the level of relation to an individual is sufficient to establish property rights, and second, in tracing the presence of such a relation when opaque AI algorithms are involved.

Ⅳ. Discussion and Conclusions

This paper aimed to revisit the European debate on the introduction of property rights in personal data in order to include the newest developments in law and data processing practices, such as the advances in Big Data. I have argued that there are clear indications that the idea of property in personal data should be given a serious consideration: the protective potential of the property rights to work as an alternative legal tool to enforce individual control over personal data, the need to acknowledge the economic value and function of personal data as a resource both for the information industry and for individuals, and finally, the de facto presence of the property-like claims of the information industry in personal data, and the need to take measures to ensure that the data-generated wealth is ~~shared fairly~~. Yet, the analysis shows that personal data does not lend itself easily as an object of individual property rights. Namely, personal data cannot be considered as concerning just an individual anymore; data processing resulting from a decision of one person will inevitably have spill-over effects on others, e.g. as a result of profiling, or as a result of the same piece of data relating to a group of people, e.g. genetic data. Therefore, true individual control over personal data and also the effective enforcement of the individual property rights in personal data are difficult if not impossible to achieve. At the same time, the reasons to create *erga omnes* protection of data grounded in the ~~data~~ economic value and ~~sharing~~ of the data-generated wealth still stand. This - in combination with the added advantages of collective and assisted consent suggested by Bygrave et al and Brownsword - points to the necessity to consider collective property rights in data and collective data management.

In addition, personal data is a difficult object of property rights when it comes to the principle of the property rights transparency. Namely, the same piece of data may behave as personal and non-personal under different circumstances: it may be more or less identifiable, and have a stronger or weaker link to a person, and the moment of transition from one state to another may pass unnoticed by the data holder and the affected individuals and groups. This makes invoking and enforcing property rights in personal data challenging. An added challenge is in identifying the boundaries of the groups or communities to whom the data pertains. The groups contributing to and affected by data, or the groups for whom the data is relevant are not always static and their boundaries shift. Taylor et al point out that it is challenging to structure accountability in the era "where almost everyone is constantly being grouped and regrouped, unaware, by data analytics".[73] Some substantial work has been done by e.g. Giseppe on 'calculated publics'[74], groups constructed by algorithms where people are influenced towards certain behaviour. What is clear is

73) L Taylor, L Floridi and B. van der Sloot, 'Conclusion: what do we know about group privacy?', in Taylor, Floridi and van der Sloot eds. n 52.

74) T Gillespie, 'The Relevance of Algorithms,' In *Media technologies: Essays on communication, materiality, and society* T. Gillespie, P. Boczkowski, & K. Foot (eds.), (Cambridge, MA: MIT Press 2014) 167-194.

that these communities have to have a sufficient degree of stability in order to have relatively stable boundaries of the data they would claim property in and manage, and to enable their participation in the data management.

Going beyond the discussion if this paper and looking into the future of the data ownership debate, I foresee that it will soon transcend the domain of personal data and shift to the domain of data generally. This is because at least two contexts of the property debate will remain valid concerning data rather than personal data, i.e. the context of acknowledging the economic value of data and its role in the economy, and the discussion on the fair distribution of the data-generated wealth.

Acknowledgements

This contribution partially reports on the results of the project "Understanding information for legal protection of people against information-induced harms" ('INFO-LEG'). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 716971). The paper reflects only the author's view and the ERC is not responsible for any use that may be made of the information it contains. The author is also thankful to the Center for Law and Public Utilities, Seoul National University, for the generous funding provided for writing this paper in part outside of the ERC project and for presenting this paper at the 2017 CeLPU International Conference 'Exploration of a New Regulatory Paradigm for the Advancement of the Fourth Industrial Revolution.'